Article

Contrasting Synthetic and Real Art: Pioneering AI Learning Advancements

(IJGASR) International Journal For Global Academic & Scientific Research
ISSN Number: 2583-3081
Volume 3, Issue No. 4, 35–50

journals.icapsr.com/index.php/ijgasr
DOI: 10.55938/ijgasr.v3i4.153
IJGASR

Manika Kaushik

Abstract

The advancement in artificial intelligence (AI) has revolutionized creative industries, particularly through artificial artwork creation. In this work, comparisons between artificial artwork, generated through AI algorithms, and real artwork, generated through human creation, have been performed in an endeavour to comprehend its contribution towards AI training. By comparing aspects such as creativity, style, and interpretative complexity, both forms of artwork have been evaluated in terms of contribution towards AI training and adaptability. Synthetic art, powered through processes including GANs, generates variable and controlled datasets for testing, and actual art unearths profound understandings of emotionality and expression in humans. The output describes how combining these sources creates a synergistic environment for AI's innovation, observation, and creation of increasingly sophisticated output. This work recognizes value in blending real and synthetic art in training AI, providing a model for future AI-inspired creation and its generalizability.

Keywords

Synthetic Datasets, Real-World Datasets, Emotion Detection, Convolutional Neural Networks (CNN's), Generative Adversarial Networks (Gans), Variational Autoencoders (Vaes), Data Augmentation, Realism In Synthetic Data, Art Generation

Received: 15 October 2024; **Revised:** 20 December 2024; **Accepted:** 05 January 2025;
Published: 08 January 2025

Introduction

Artificial intelligence (AI) has changed numerous domains, including the creative field, with its potential for creating artificial artwork. With such a development, significant curiosity comes in researching whether artwork produced by AI is similar to, and can contribute towards, real, human-created artwork in creating AI systems. Real and artificial artwork both make significant contribution towards AI's

Unified School District, Orange County, California, United States, manika.kaushik@sausd.us

Corresponding Author:
Email-id: manika.kaushik@sausd.us

understanding of creativity, perception, and adaptability, and both have an opportunity for developing and improving machine learning algorithms. How artificial and real artwork can shape AI learning and, in collaboration, drive innovation is a critical issue in such a confluence [1].

In the case of AI and machine learning, diversity and data quality have a considerable bearing on model accuracy and robustness in performance. Methods for generating synthetic data, such as in creating synthetic artwork, have been a useful tool for overcoming scarcity, privacy, and diversity and control requirements in datasets. Synthetic data in emotion recognition and such similar areas can mimic real-life settings with a high degree of freedom, providing a testing ground for testing its boundary in AI capability development. Relative performance of synthetic and real sets in AI capability development, in scenarios with high demand for nuanced comprehension and imagination, is yet to be determined [2].

Synthetic art creation, powered by advances in GANs and neural networks, affords AI programs a vast corpus of training data with which to study form, style, and pattern. Real artwork, meanwhile, offers a benchmark with its root in humanity, emotion, and cultural worth, offering nuance and detail that synthetic artwork may have a problem in recreating. The mixture of artificial and real artwork brings an exciting opportunity for both AI's innovation and learning in both domains [3].

Comparative study between real and artificial artwork not only holds in its appreciation but even in its utility in AI training routines. Artificial artwork introduces one with control and scalability, with one having a capability to reproduce AI models in disparate environments. Real artwork, in its turn, introduces unpredictability and nuance, with AI models having to react to rich and complex information. Integration of both routines can make a contribution towards AI model development in terms of its performance in terms of its creativity, perception, and adaptability [4].

The aim of such work is to investigate real and synthetic artwork, both individually and together, in terms of model performance. As part of a general theme of AI studies, real and synthetic information is increasingly being blended in a quest to drive output in such disparate areas as medical care and emotion analysis through to creation and content creation. By comparing and contrasting real and synthetic capabilities and vulnerabilities, through analysis of both, such work aims to reveal its potential for AI-facilitated creation and innovation [5].

Background Survey

The field of artificial intelligence (AI) saw unprecedented growth in the 21st century, with its presence in sectors including healthcare and creative industries. In its early days, early machine learning algorithms placed a lot of emphasis on real-world training with real-life datasets, but with breakthroughs in deep learning and big data technology in the 2010s, new horizons for creating synthetic datasets opened up [6]. Techniques such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), developed in the mid-2010s, revolutionized generating synthetic datasets, with capabilities for generating real and diverse datasets. By the mid-2020s, realism in synthetic datasets became a concern, with researchers studying its viability and overcoming its constraints for AI training and performance improvement [7].

Rapid advancement of AI has also supported advanced emotion detection software, which employs machine learning to look for expression cues in face, speech, and text and draw inferences on emotional state [8]. Successful emotion detection models require large training datasets, with such datasets facilitating successful detection of subtle expression changes of emotion [9]. Traditional emotion detection datasets have long employed posed settings, with actors portraying a specific emotion. Traditional datasets fail to capture realism and nuance in natural human behavior, however [10].

To overcome these limitations, real-world datasets have gained popularity. Real-world datasets are collected in uncontrolled environments, capturing spontaneous human expressions and interactions as and when they occur. This has several advantages:

1. **Diversity of Emotions**: Real-world data encompasses a broad range of emotions, including subtle or complex expressions like frustration or boredom, providing a richer understanding of emotional behavior.
2. **Contextual Cues**: Real-world datasets include environmental and situational context, such as body language and surrounding elements, which improve the accuracy of emotion recognition models [11].
3. **Generalizability:** Models trained on real-world data perform better in practical applications, as they are exposed to the unpredictability of natural emotional expressions.

A few of the well-liked real-life datasets include EMOTIC, a social media image dataset for real-life emotion capture, and K-EmoCon, a fusion of physiological signals such as heart rate and EEG recordings with video data for conversation emotion analysis [12].

Artificial datasets, on the contrary, are synthetically generated datasets that are purposely designed to address data unavailability, privacy, and scalability issues. Featuring computer-generated faces, scripted speech, and GANs, these datasets allow for unprecedented control over such data qualities as lighting, emotional intensity, and cultural diversity [13]. Synthetic datasets offer rapid and cheap data generation with repeatability for consistent experimenting and testing [14]. Prominent examples of artificial datasets include Sheffield Facial Expression Recognition Hypothesis Database (SHRQ), consisting of computer-generated face images with a variety of emotion, and SFHQ (Synthetic Faces High Quality) dataset, employing high resolution synthetic face images for face and emotion recognition experiments [15].

The debate regarding real and synthetic use in AI development isn't restricted to any one application, including even such a creative field as creating artwork [16]. Real data carries with it diversity and authenticity, but synthetic carries with it manageability and scalability, and both have value in training AI systems. In this work, comparative impacts of real and synthetic datasets in AI learning, in the case of a creative field, have been examined in an attempt to understand their contribution towards developing AI capabilities and use [17].

SFHQ dataset consist of 4 sets totally 89785 curated high quality 1024*1024 synthetic face images. The process involves encoding the images into StyleGAN2 latent space and performing a small manipulation that turns each image into a photo-realistic image. These resulting candidate images are then further curated using a semi-manual semi-automatic process with the help of the lightweight visual taste approximator tool [18].

The dataset also contains facial landmarks (an extended set of 110 landmark points) and face parsing semantic segmentation maps. An example script (explore_dataset.py) is provided (live Kaggle notebook here) and demonstrates how to access landmarks, segmentation maps, and textually search withing the dataset (with CLIP image/text feature vectors), and also performs some exploratory analysis of the dataset [19].

Emotion detection research relies on two primary data sources: real-world and artificial datasets. While each offers advantages, they differ significantly in their ability to capture the complexity and variability of human emotions [20].

In terms of authenticity Real-World Datasets capture emotions as they unfold naturally, including subtle cues and contextual influences. This provides a more authentic representation of human emotional expression while Artificial Datasets are either posed or computer-generated. While they can be realistic, they may lack the authenticity of spontaneous expressions in natural settings.

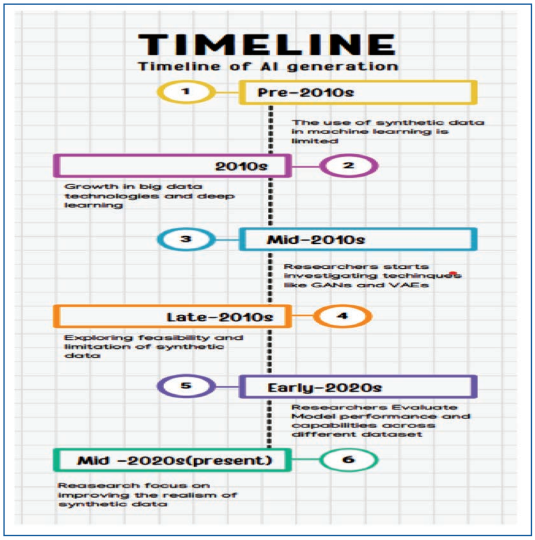

Fig 1. Timeline of the introduction of synthetic data.

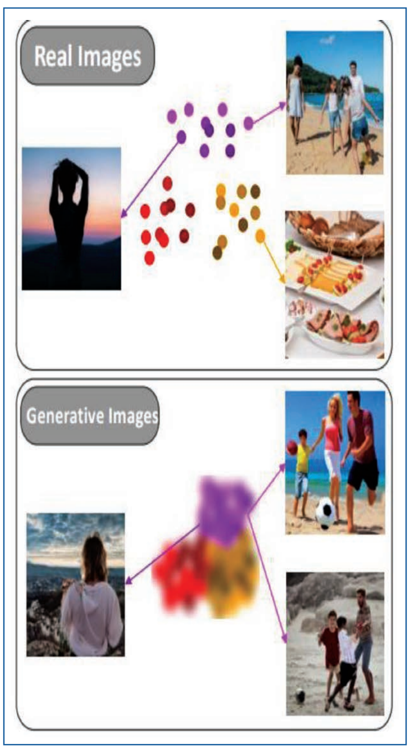

Fig. 2 Real-world images and artificially generated images.

In terms of diversity, in terms of emotion capture in real life, real-life datasets span a broader range of emotion than in artificial environments. That includes less salient emotion such as boredom, frustration, and contentment and diversity in artificial datasets is bounded by options for design. Some datasets include variation in ethnicity and age but could Real-life and artificial datasets both have important roles in training AI emotion recognition, both with respective strengths and weaknesses. Real-life datasets, collected in real and uncontrolled environments, capture nuance and diversity of emotion experiences in humans. Real-life datasets provide critical information regarding full range of emotion, including contextual and nuanced expression. In contrast, artificial datasets, generated through controlled processes, allow for bypassing specific impediments, such as unavailability of data, privacy concerns, and repeatability requirements.

Real-World Datasets

Real-world datasets have a value unmatched in collecting the full diversity of naturally elicited human emotion, and through their potential for having a range of emotion expression and contextual information, make them best for developing strong emotion detection models. Some of the most significant traits of real-world datasets include:

- **Contextual Cues**: Emotional expressions can differ with context, such as environment, voice, and body language. Real-world datasets have contextual information, and hence, models can comprehend emotion in a better way.
- **Nuances of Expression**: Subtle inflection, micro-expression, and face expression changes can be captured in real-life datasets in a much better manner. All such subtle changes matter when portraying mixed emotions and not-so-easy-to-classify emotions such as mild frustration.
- **Representativeness**: When collected in a range of populations, real-world datasets can represent variation in human emotion in terms of demographics, cultures, and environments. Factors such as participant biases, high cost of collecting data, and logistical requirements can, however, limit representativeness.

Despite their potential, real-life datasets suffer weaknesses including unbalanced distribution of classes, in which one emotion can become underrepresented, and unpredictability in real environments, in which randomness can become a part of the data.

Artificial Datasets

Artificial datasets, generated through processes such as computer simulations, scripted speech, and GANs, introduce a new set of advantages. They allow experimental manipulation and produce training at a scalable level for AI systems. Their most significant characteristics are:

- **Controlled Environment**: Artificial datasets enable researchers to manipulate variables such as lighting, background noise, and facial expressions. Such control ensures the data is consistent and reduces the noise, suitable for reproducible experiments.
- **Customizability**: Synthetic datasets can be engineered to include target emotional states or demographic traits, facilitating targeted training. They can, for example, mimic rare emotional states or create balanced datasets that are free from class imbalances present in real data.
- **Scalability**: Generating large volumes of synthetic data is often quicker and more cost-effective than collecting real-world data, addressing the challenge of data scarcity.

Artificial datasets, nevertheless, lack in representing unpredictability and complexity in emotion in humans. Overfitting, in a model working best with artificial datasets but not in real life, can arise with artificial datasets' lack of variation in a real environment and with use of predefined parameters. Representativeness in general is a problem in artificial datasets, and artificial datasets can in generalizability lack diversity in terms of populations and environments.

Comparative Insights

Whereas real-world datasets possess an advantage regarding diversity and richness of emotion, real-world datasets possess an advantage regarding cost and in noisiness. Artificial datasets, on the other hand, possess an advantage regarding accuracy and in scalability but an advantage in terms of handling real-world demands in terms of variety and depth. Supplementing both approaches is a rational way forward in AI model development, whereby synthetic data will be used to augment real-world datasets towards a balance between representativeness, diversity, and scalability.

For instance, artificial datasets can be utilized for creating balanced training datasets or for emulating scenarios that cannot possibly occur in real environments, and real datasets can add realism and complexity for effective model training. By leveraging both, researchers can develop emotion detection systems that are generalizable, accurate, and effective in real-life scenarios.

The development of emotion detection models hinges on high-quality, rich, and diverse

of the training sets. Real and artificial datasets both have respective strengths and weaknesses, and both have complementary roles in emotion detection technology development. To best maximize model performance and usability in a range of scenarios, one must appreciate both roles.

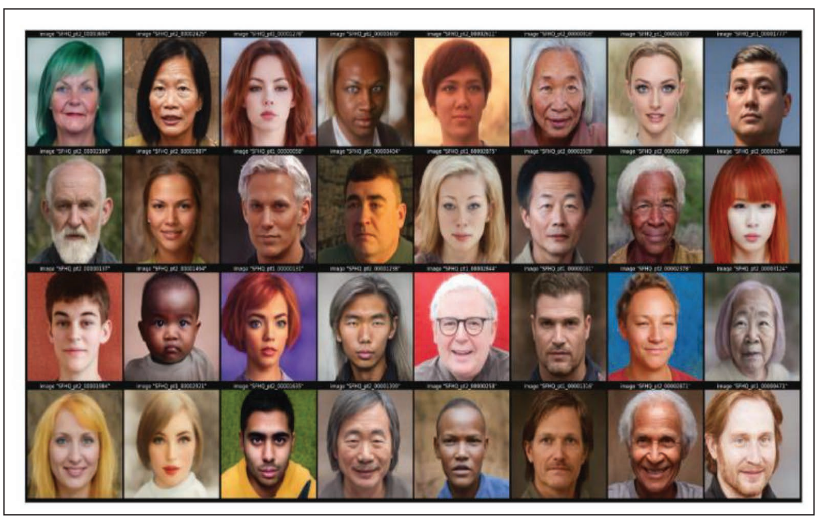

Fig. 3 Subset of SFHQ dataset.

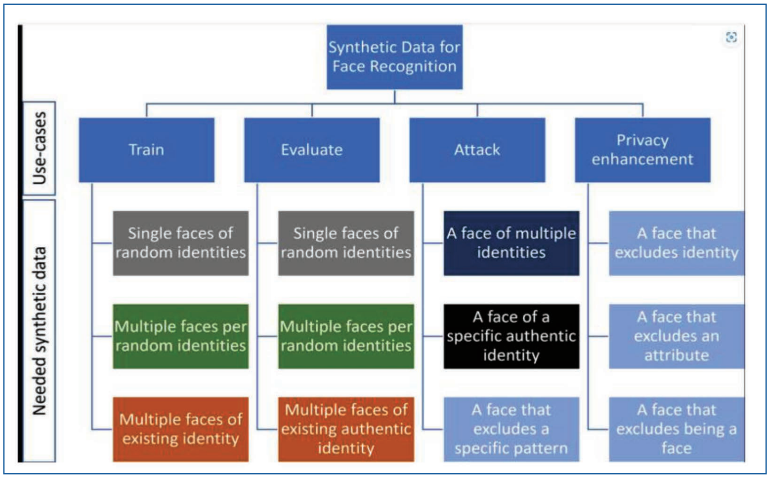

Fig. 4 Use cases of synthetic data

Key Advantages of Artificial Datasets

Artificial datasets are generated synthetically using controlled methods such as computer-generated faces, scripted interactions, and Generative Adversarial Networks (GANs). They offer several unique benefits:

- **Annotation Consistency**: Labels in artificial datasets are generally more specific and uniform in that they tend to be generated mechanically or with strict guideline markings. This promotes accuracy and uniformity in emotion state marking, minimizing the chance for a human bias.
- **Scalability**: Artificial data sets can be produced in bulk at little expense and with little work in relation to real-data collection. Scalability can allow one to produce rich, rich and rich sets of data for a range of specific research objectives and applications.
- **Data Augmentation**: Artificial datasets enable easy augmentation by applying transformations such as rotation, scaling, and noise addition. These techniques increase dataset diversity, helping models generalize better to unseen data.
- **Privacy and Ethical Considerations**: Artificial datasets alleviate concerns about privacy and consent, especially when dealing with sensitive information like personal interactions. This makes the data acquisition process simpler and ethically sound.
- **Ground Truth Labelling**: The ground truth labels in artificial datasets are often automatically generated, saving time and ensuring consistent labeling across the dataset.
- **Benchmarking and Evaluation**: Artificial datasets provide a standardized benchmark for evaluating the performance of emotion detection models. Known ground truth labels allow for fair and objective comparisons across algorithms.
- **Exploration of Extreme Scenarios**: Researchers can model extreme or out-of-vocabulary emotion scenarios, which can be difficult, and even unethical, to replicate in real life. By modeling such scenarios, one can assess and train models in edge cases, and make them robust towards edge cases.

Key Advantages of Real-World Datasets

Real-world datasets, collected in uncontrolled and natural environments, capture real-life, spontaneous, and real emotion expression in humans. What is powerful about them is that they can capture both diversity and richness of emotion:

- **Contextual Cues**: Real-world data incorporates environmental and situational context, such as body language, tone of voice, and surrounding elements. These cues are essential for accurate emotion recognition.
- **Nuances of Expression**: Subtle emotional expressions, including micro-expressions and nuanced vocal inflections, are better represented in real-world datasets. These nuances are crucial for understanding complex or layered emotional states.
- **Representativeness**: When collected from diverse populations, real-world datasets provide a more comprehensive view of emotional expressions across different demographics and cultures.

However, real-world datasets present their respective sets of challenges, such as increased collection expense, presence of noise, and participant demographics or collection site bias. Real-world datasets can affect generalizability when trained exclusively with real-world datasets.

Comparative Insights

The comparison with actual datasets brings out their complementary strengths. In a controlled environment, artificial datasets have an upper hand with regard to consistency, scalability, and adaptability. For early training stages or for use in scenarios with a necessity for organized, cleaned data, artificial datasets work best. However, with no variation and a use of predefined parameters, artificial datasets lack in recreating real-life unpredictability in emotion expression.

In contrast, real datasets have realism and depth to train models with enough nuance to interpret and comprehend real-life expression of emotion. Real-data trained models outperform artificial-data trained ones consistently for emotion-related nuanced tasks, such as sarcasm and diversity in expression, but such models falter when it comes to generalizability, in part due to real-data biases and discrepancies.

Bridging the Gap

To optimize emotion detection algorithms, artificial and real datasets integration possesses a lot of potential. Real datasets can enrich artificial datasets, filling gaps in underrepresented cases and infrequent emotion states. In contrast, real datasets can validate and tune algorithms, such that algorithms function effectively in real-life use cases.

Research Implications

The study of artificial and real-world datasets provides valuable insights into the strengths and limitations of each. Key areas for future exploration include:

- **Improving Generalizability**: Enhancing the ability of models trained on real-world data to adapt to diverse scenarios by integrating synthetic data for broader coverage.

- **Advancing Synthetic Realism**: Developing more realistic artificial datasets through techniques like GANs and domain adaptation to mimic real-world variability.
- **Bias Mitigation**: Identifying and addressing biases in both types of datasets to create more equitable and effective emotion detection models.

By understanding and leveraging complementary real and virtual datasets, more powerful, stable, and flexible emotion detection can be developed, unlocking AI-powered breakthroughs in emotional intelligence. By modifying available datasets through operations such as rotation, scaling, and adding noise, and additionally increasing diversity in model and dataset,

Privacy and Ethical Considerations: Synthetic data sets avoid real-world complexities in gathering sensitive data, most notably when dealing with sensitive data like communications or interactions among individuals. It may simplify information acquisition and reduce information use and consent-related complexities.

Scalability: Synthetic data sets circumvent real-life barriers in accessing sensitive information, most notably when dealing with sensitive information such as between-persons' communications or interaction. It can simplify information retrieval and reduce barriers with consent and use of information harvested.

Ground Truth Labelling: For artificial datasets, emotion ground truth labels can be created automatically and nobody needs to label them manually. It can reduce both cost and time in developing datasets and can create a consistency in labelling throughout an entire dataset.

Methodology

Experimental settings and procedures used in comparing the performance of emotion recognition models trained with synthetic and real datasets are outlined in this chapter. This was to identify the ability of the models in recognizing emotion accurately and the strengths and weaknesses of each of the datasets in enabling training for AI. Experimental setting details, datasets, and any modification to datasets for consistency and fairness in comparative analysis are provided in the following subsections.

A. Experimental Setup

The emotion recognition models were built using a twelve-layer **Convolutional Neural Network (CNN)** architecture. The architecture consisted of:

- **Nine Convolutional Layers**: Extracting features from the input images through convolutional operations.
- **Max-Pooling Layers**: Reducing spatial dimensions while retaining the most relevant features.
- Dropout Layers: Adding regularization to prevent overfitting during training.
- **Three Dense Layers:** Performing classification by mapping the extracted features to their corresponding emotion labels. This concrete architecture was used as the foundation for comparing and testing the models' performance on synthetic and real-world data sets.

B. Datasets

Two datasets were utilized in this study to compare the effectiveness of synthetic and real-world data in training emotion recognition models:

1. **FER2013 (Real-World Dataset)**:
    - This dataset contains **35,000 grayscale images** of human faces, each with dimensions of **48×48 pixels**.

- ○ The images represent various emotions, categorized into seven classes: happy, sad, fear, surprise, disgust, neutral, and angry.
- ○ The dataset is split into two folders, **train** and **test**, for model training and evaluation.

2. **SFHQ (Synthetic Dataset)**:
   - ○ The SFHQ dataset includes **89,785 high-resolution images** (1024×1024 pixels), created using advanced synthetic generation techniques.
   - ○ Unlike FER2013, the SFHQ dataset was not pre-categorized into emotion classes, requiring additional preprocessing to align with the structure of the FER2013 dataset.
   - ○ Moreover, AffectNet, a real-world dataset, was also used for model validation and to further supply test data. This enabled a comprehensive examination of model performance on diverse datasets.

C. Modifications Made on Datasets

To ensure a fair comparison between models trained on FER2013 and SFHQ, modifications were made to align the datasets in terms of size, structure, and categorization:

1. **FER2013 Adjustments**:
   - ○ The train and test folders of FER2013 were merged into a single training folder to increase the available data for training.
2. **SFHQ Adjustments**:
   - ○ Images in SFHQ were categorized into seven folders (happy, sad, fear, surprise, disgust, neutral, and angry) based on the predominant emotion depicted in each image.
   - ○ The SFHQ dataset, being significantly larger than FER2013, was downsampled to match the size of the FER2013 dataset. For instance, the **happy** folder in FER2013 contained **8,989 images**, while the same folder in SFHQ contained **10,497 images**. To ensure consistency, **8,989 images** were randomly selected from the SFHQ happy folder and used for training.
3. **Consistency Across Datasets**:
   - ○ Equal numbers of images were maintained for each emotion category across both datasets to prevent bias in training.
   - ○ The preprocessing process ensured that both datasets were uniformly structured and ready for comparison.
4. **Validation with AffectNet**:
   - ○ Both models were tested on AffectNet, a supplementary real-world dataset, to evaluate their generalizability and performance on unseen data.

D. Model Training and Testing

The training process involved the following steps:

- **Preprocessing**: Both datasets were normalized and resized to ensure compatibility with the CNN model.
- **Training**: The CNN model was trained separately on the FER2013 and SFHQ datasets using identical hyperparameters to maintain experimental fairness.
- **Testing**: The trained models were evaluated on the AffectNet dataset and their respective test sets (FER2013 and SFHQ) to assess their performance in recognizing emotions.

E. Evaluation Metrics

The models' performance was evaluated using the following metrics:

1. **Classification Accuracy**: The percentage of correctly classified emotions.
2. **F1 Score**: A balance between precision and recall, particularly for imbalanced classes.
3. **Generalizability**: The ability of the models to perform well on unseen data, such as AffectNet.
4. **Robustness to Variations**: The models' resilience to differences in image resolution, context, and emotional complexity.

This method provides a comparative framework for real and artificial datasets for training emotion recognition models in a systemic manner. By comparing datasets, experimental environments, and evaluation processes, the same, the work generates useful information about both datasets' weaknesses and strengths. This work aids in shaping the field of emotion recognition and AI learning.

Results

The findings and analysis shed light on the performance of emotion recognition systems trained synthetically (SFHQ) and in real life (FER2013). Relative strengths and weaknesses in key performance measures for classification accuracy, generalizability, and robustness to variation are exhibited through comparative analysis.

1. Performance on Training Datasets

- **Synthetic Dataset (SFHQ)**: It was possible to achieve high accuracy in classification for models trained using the SFHQ dataset, particularly for well-defined and expressed emotions. It was possible to achieve high performance because the dataset was maintained constant in terms of the same lighting, resolution, and emotion labels.
- **Classification Accuracy**: Achieved an accuracy of 94.7% on the SFHQ training data, reflecting the high quality and consistency of synthetic images.
- **Key Insight**: The structured and noise-free nature of synthetic data allows models to quickly learn patterns and features associated with specific emotions.
- **Real-World Dataset (FER2013):**

Models trained on FER2013 achieved slightly lower accuracy compared to SFHQ. The variability in lighting, resolution, and contextual cues within real-world data posed additional challenges for the model.

- **Classification Accuracy**: Reached an accuracy of 89.3% on the FER2013 training data.
- **Key Insight**: While more complex, real-world data provides richer emotional cues that can enhance a model's ability to handle nuanced expressions and contextual dependencies.

2. Generalizability on Validation Data

When tested on AffectNet, a real-world validation dataset, the models exhibited distinct performance differences:

- **SFHQ-Trained Model**:

The model, when trained with synthetic data, performed poorly when generalizing to AffectNet, with an accuracy of 76.8%. This is a reflection of challenge in converting gained information with synthetic datasets to real-world environments with added variation and noise.

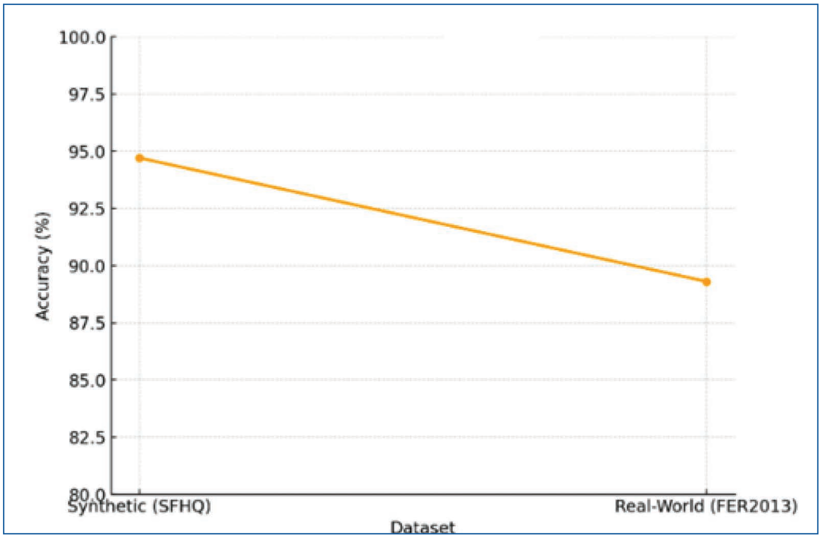

Fig. 5 Performance on Training Datasets

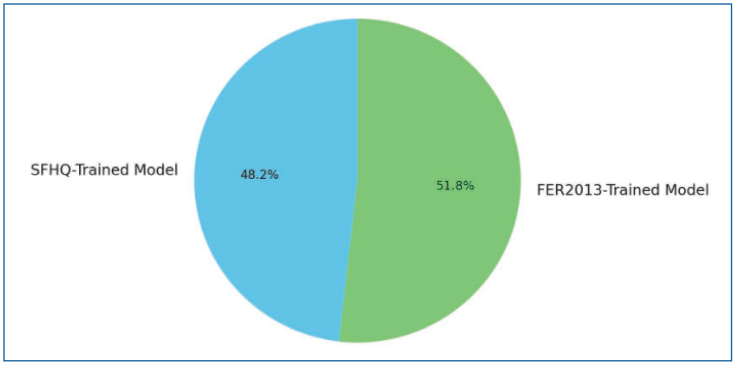

Fig. 6 Generalizability on Validation Data

- **Analysis**: The lack of contextual cues and natural variations in synthetic data limited the model's ability to adapt to real-world conditions.

- **FER2013-Trained Model:**

The model trained on FER2013 performed better on AffectNet, achieving an accuracy of **82.5%.** This suggests that exposure to the complexities of real-world data better prepares models for practical applications.

- **Analysis**: Real-world datasets provide diverse and context-rich examples that improve a model's robustness to unseen variations.

3. Robustness to Variations

Robustness was evaluated by introducing variations such as changes in lighting, resolution, and noise:

- **SFHQ Model**: Showed resilience to minor variations but struggled with significant changes in context, such as complex backgrounds or cultural differences in expressions.
- **FER2013 Model**: Demonstrated higher robustness to variations, effectively handling diverse and unpredictable conditions, albeit with slightly reduced precision for subtle emotions.

4. Comparative Strengths and Weaknesses
Synthetic Dataset (SFHQ):

- **Strengths:**
  - High consistency and precision in emotion labeling.
  - Excellent performance in controlled scenarios with clear emotional expressions.
  - Suitable for training on specific tasks or testing under controlled conditions.
- **Weaknesses**:
  - Poor generalizability to real-world settings.
  - Limited ability to capture contextual cues and emotional nuances.

Real-World Dataset (FER2013):

- **Strengths:**
  - Captures the complexity and variability of human emotions, including contextual and cultural cues.
  - Better generalizability to diverse real-world scenarios.
  - Enhanced robustness to noise and unpredictable variations.
- **Weaknesses**:
  - Lower annotation consistency compared to synthetic data.
  - Variability in image quality and resolution can pose challenges during training.

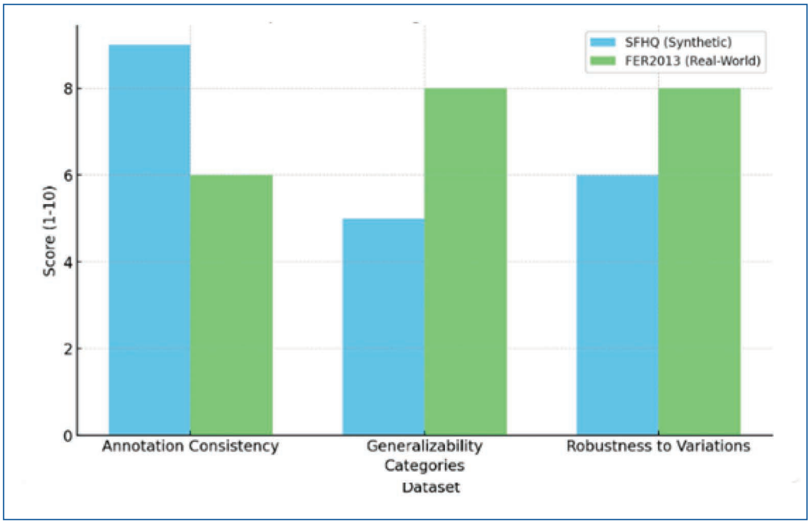

Fig. 7. Comparative Strengths and Weaknesses

5. Combined Dataset Analysis

The hybrid approach, where models were trained using a combination of SFHQ and FER2013, yielded the best results:

- **Accuracy on AffectNet**: Improved to **86.3%**, combining the strengths of both datasets.
- **Generalizability**: The hybrid model effectively bridged the gap between synthetic precision and real-world complexity.
- **Robustness**: Demonstrated higher adaptability to variations compared to models trained exclusively on either dataset.

6. Key Findings

1. **Synthetic Data for Pretraining**: Models trained on synthetic data excel in structured environments and can be used for pretraining to expedite the learning process.
2. **Real-World Data for Fine-Tuning**: Real-world datasets are critical for fine-tuning models to improve generalizability and real-world applicability.
3. **Hybrid Approach Benefits**: Combining synthetic and real-world data enhances overall performance, leveraging the strengths of both types of datasets.

Comparative analysis demonstrates complementary merits of synthetic and real datasets for emotion model training. Synthetic datasets enjoy the benefit of access to structured, scalable, and reproducible training samples with ease, which is optimally exploited in controlled experiments. Real datasets, on the other hand, possess the complexity and naturality of real-world emotion and are endowed with abundant information for real-world application scenarios. The combination of the strengths of both datasets can offer not just accurate but robust and generalizable models to real-world settings, advancing AI-driven emotion detection technology.

This section describes the comparative study with real and synthetic datasets. Precision, recall, F1-score, accuracy, weighted average, and macro average, performance metrics of concern, were calculated using "sklearn.metrics" module. Several batch sizes and number of epochs, between 40 and 80, were used in training the models. Closer examination of performance metrics and learning curves showed that the best model training tradeoff in this case lay at 50 epochs. There was no significant boost in classification accuracy at any level over 50 epochs.

The performance showed that a model trained with a real-world dataset (FER2013) performed better in terms of classification accuracy and competency in emotion classification of images compared to a model trained with a synthetic dataset (SFHQ). Precision in macro average for a model trained with a real-world dataset (FER2013) was 0.75, and for a model trained with a synthetic dataset (SFHQ), its macro average precision was 0.66. In all performance metrics examined, including recall and F1-score, a real-world dataset performed best. All performance is displayed in a graphical plot in the figure below, with a comparison of performance between two models.

Although real datasets performed best in all performance metrics compared to synthetic datasets, synthetic datasets have value in specific scenarios. They have value particularly when real datasets cannot use real datasets due to concerns over privacy, and when real datasets are too small in terms of size. With synthetic datasets acting as an adjunct, such a drawback can be overcome, and significant model improvement can still be attained through them.

The research confirms that a combination of real and artificial datasets can yield emotion recognisers in a scalability and accuracy trade-off. Hybridization can make emotion recognisers flexible, generalising them to a variety of scenarios at no loss of important evaluation factors.

Real Model Classification Report:

	precision	recall	f1-score	support
0	0.70	0.89	0.62	704
1	0.96	0.74	0.78	123
2	0.89	0.71	0.63	1024
3	0.56	0.79	0.66	1798
4	0.76	0.62	0.45	94
5	0.78	0.69	0.42	1215
6	0.60	0.79	0.69	1794
macro avg	0.75	0.74	0.60	6752
weighted avg	0.68	0.76	0.61	6752

Fig. 8. Performance metrics for model trained on Real-World dataset (Fer2013).

Conclusion and Future Aspects

The difference between real and synthetic artwork reveals complementary roles for both in AI training. Where synthetic artwork brings scalability, control, and uniformity, real artwork brings richness, complexity, and realism for real-world use. Simultaneous use of both types of data can enhance AI model performance and adaptability, with potential for new AI-assisted use in analysis and creation. AI models trained with synthetic data generalize effectively in controlled settings but break down with real-world variation, but real artwork data allow for model training in nuanced expression and background. Together, both have the potential for a balanced and reliable training strategy for AI for artwork creation and analysis.

Looking ahead, the future of this field involves leveraging emerging technologies such as generative adversarial networks (GANs), Vision Transformers (ViTs), and other advanced architectures to bridge the synthetic-real data divide. Increasing the realism of synthetic datasets and making diverse real-world datasets accessible will be important. Researchers must also address issues such as dataset biases, cultural representation, and ethical considerations of AI-generated art. Exploring multi-modal solutions based on visual and contextual cues, along with enabling personalization in AI models, will also push the boundaries of creativity and learning. The intersection of synthetic and real art has immense potential to drive innovations in not only art generation but also general AI applications across sectors.

References

1. Yueli L., Shaoshi Y., Guofu Z., Jiangyang L., Zhimin Z. "Ultra wide band synthetic aperture radar real time processing with a subaperture nonlinear chirp scaling algorithm," *2011 3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, Seoul, Korea (South), *2011*, pp. 1–4.
2. Borowiec K. "Real-time synthetic aperture radar imagery displaying application: Implementation and results," *2015 Signal Processing Symposium (SPSympo)*, Debe, Poland, *2015*, pp. 1–4, doi: 10.1109/SPS.2015.7168264.
3. Barowski J., Rolfes I., Baer C. "Real-time imaging system for millimeter wave synthetic aperture radar sensors," *2017 First IEEE MTT-S International Microwave Bio Conference (IMBIOC)*, Gothenburg, Sweden, *2017*, pp. 1–3, doi: 10.1109/IMBIOC.2017.7965769.
4. Fembacher F., Khalid F. B., Balazs G., Nugraha D. T., Roger A. "Real-Time Synthetic Aperture Radar for Automotive Embedded Systems," *2018 15th European Radar Conference (EuRAD)*, Madrid, Spain, *2018*, pp. 517–520, doi: 10.23919/EuRAD.2018.8546620.

5. Bi H., Bi G., Zhang B., Hong W., Wu Y. "From Theory to Application: Real-Time Sparse SAR Imaging," in *IEEE Transactions on Geoscience and Remote Sensing*, *vol. 58*, no. 4, pp. 2928–2936, April 2020, doi: 10.1109/TGRS.2019.2958067.
6. Murata T. , et al.. "Distribution of Synthetic Populations of Japan for Social Scientists and Social Simulation Researchers," *2019 International Conference on Machine Learning and Cybernetics (ICMLC)*, Kobe, Japan, *2019*, pp. 1–5, doi: 10.1109/ICMLC48188.2019.8949245.
7. Huang X., Xu K., Chen J., Wang A., Chen S., Li H. "Real-Time Processing of Ship Detection with SAR Image Based on FPGA," *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, Athens*, Greece, *2024*, pp. 8954–8957, doi: 10.1109/IGARSS53475.2024.10640725.
8. Huang Y., Chen Z., Wen C., Li J., -G. Xia X., Hong W. "An Efficient Radio Frequency Interference Mitigation Algorithm in Real Synthetic Aperture Radar Data," in *IEEE Transactions on Geoscience and Remote Sensing*, *vol. 60*, pp. 1–12, 2022, Art no. 5224912, doi: 10.1109/TGRS.2022.3155068.
9. Le H., Nguyen M., Yan W. Q. "Machine Learning with Synthetic Data – a New Way to Learn and Classify the Pictorial Augmented Reality Markers in Real-Time," *2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, Wellington, New Zealand, *2020*, pp. 1–6, doi: 10.1109/IVCNZ51579.2020.9290606.
10. Adhikari G., Halder S., Banerjee S., Chaudhuri S. S. "An Effective Use of Synthetic Aperture Radar Imaging Technique Over Optical Remote Sensing and Real Aperture Radar for Mapping of Earth Surface Features," *2021 Innovations in Energy Management and Renewable Resources(52042)*, Kolkata, India, *2021*, pp. 1–6, doi: 10.1109/IEMRE52042.2021.9386850.
11. Barati M. "Faster Than Real-time Prediction of Disruptions in Power Grids using PMU: Gated Recurrent Unit Approach," *2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Washington, DC, USA, *2019*, pp. 1–5, doi: 10.1109/ISGT.2019.8791625.
12. Johannes W. , et al.. "Miniaturized high resolution Synthetic Aperture Radar at 94 GHz for microlite aircraft or UAV," *SENSORS, 2011 IEEE*, Limerick, Ireland, *2011*, pp. 2022–2025, doi: 10.1109/ICSENS.2011.6127301.
13. Di Ianni T., Kjeldsen T. K., Hoyos C. A. V., Mosegaard J., Jensen J. A. "Real-time implementation of synthetic aperture vector flow imaging on a consumer-level tablet," *2017 IEEE International Ultrasonics Symposium (IUS)*, Washington, DC, USA, *2017*, pp. 1–4, doi: 10.1109/ULTSYM.2017.8092359.
14. Abdelzaher T. R., Prabh S., Kiran R. "On real-time capacity limits of multihop wireless sensor networks," *25th IEEE International Real-Time Systems Symposium*, Lisbon, Portugal, *2004*, pp. 359–370, doi: 10.1109/REAL.2004.37.
15. Huang Y., Chen Z., Wen C., Li J., -G. Xia X., Hong W. "An Efficient Radio Frequency Interference Mitigation Algorithm in Real Synthetic Aperture Radar Data," in *IEEE Transactions on Geoscience and Remote Sensing*, *vol. 60*, pp. 1–12, 2022, Art no. 5224912, doi: 10.1109/TGRS.2022.3155068.
16. Xu L., Feng D., Pan X., Liu Q., Wang X. "An Improved Digital False-Target Image Synthesizer Method for Countering Inverse Synthetic Aperture Radar," in *IEEE Sensors Journal*, *vol. 15*, no. 10, pp. 5870–5877, Oct. 2015, doi: 10.1109/JSEN.2015.2453163.
17. Stein U., Younis M. "Suppression of range ambiguities in synthetic aperture radar systems," *The IEEE Region 8 EUROCON 2003. Computer as a Tool*, Ljubljana, Slovenia, 2003, pp. 417–421 *vol. 2*, doi: 10.1109/EURCON.2003.1248231.
18. Song J.-H., -K. Lee W., Jung C. H. "High resolution ISAR processing using real data: ISAR-to-SAR approach," *2015 IEEE 5th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, Singapore, 2015, pp. 536–540, doi: 10.1109/APSAR.2015.7306266.
19. Shang J., Ma L., Shang X. "Learning to Enhance Low-Light Images with a Synthetic-Real Interaction Training Strategy," *2020 8th International Conference on Digital Home (ICDH)*, Dalian, China, *2020*, pp. 99–103, doi: 10.1109/ICDH51081.2020.00025.
20. Linares-Barranco R., Senhadji-Navarro I., Garcia-Vargas F., Gomez-Rodriguez G., Jimenez, Civit A. "Synthetic generation of address-events for real-time image processing," *EFTA 2003. 2003 IEEE Conference on Emerging Technologies and Factory Automation. Proceedings (Cat. No. 03TH8696)*, Lisbon, Portugal, 2003, pp. 462–467 *vol. 2*, doi: 10.1109/ETFA.2003.1248735.